Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1. (Currently amended) A ~~computer-implemented~~ method performed by a server device, of identifying whether a sequence of terms is a semantic unit, the method comprising:

receiving, by a communication interface or an input device of the server device, the sequence of terms in a memory;

calculating, by a processor of the server device, a first value representing a coherence of terms in the sequence;

calculating, by the processor, a second value representing variation of context in which the sequence occurs;

comparing, by the processor, the first value to a first threshold and the second value to a second threshold;

~~determining~~ identifying, by the processor, that ~~whether~~ the sequence is a semantic unit based at least in part on the first value satisfying the first threshold and the second value satisfying the second threshold ~~values~~ ; and

outputting, by the communication interface or an output device of the server device, an indication ~~of whether~~ that the sequence is a semantic unit based on identifying that the sequence is a semantic unit ~~for use in a processor~~.

2. (Currently amended) The method of claim 1, ~~wherein~~ where the coherence of the terms in the sequence is calculated relative to a collection of documents.

3. (Currently amended) The method of claim 2, ~~wherein~~ where the coherence of the terms in the sequence is calculated as a likelihood ratio that defines a probability of the sequence occurring in the collection of documents relative to parts of the sequence occurring.

4. (Currently amended) The method of claim 2, ~~wherein~~ where the coherence of the terms in the sequence is calculated as:

$$LR(A,B) = \frac{L(f(B),N)}{L(f(AB),f(A)) \cdot L(f(\sim AB),f(\sim A))},$$

where $f(A)$ defines a number of occurrences of term A in the collection of documents, $f(\sim A)$ defines a number of occurrences of a term other than term A in the collection of documents, $f(B)$ defines a number of occurrences of term B in the collection of documents, N defines a total number of events in the collection of documents, $f(AB)$ defines a number of times term A is followed by term B in the collection of documents, and $f(\sim AB)$ is a number of times a term other than A is followed by term B in the collection of documents, ~~wherein~~ where

$$L(k,n) = \left(\frac{k}{n}\right)^{k} \cdot \left(1 - \frac{k}{n}\right)^{(n-k)} .$$

5. (Canceled)

6. (Currently amended) The method of claim [[5]] 1, ~~wherein~~ where the first threshold is defined as: $f(AB) > \frac{f(A) \cdot f(B)}{N}$, where $f(A)$ defines a number of occurrences of term A in the collection of documents, $f(B)$ defines a number of occurrences of term B in the collection of documents, N defines a total number of events in the collection of documents, and $f(AB)$ defines a number of times term A is followed by term B in the collection of documents.

7. (Currently amended) The method of claim 1, ~~wherein~~ where the variation of context in which the sequence occurs is calculated relative to a collection of documents.

8. (Currently amended) The method of claim 7, ~~wherein~~ where the variation of context in which the sequence occurs is calculated as a measure of entropy of the context of the sequence.

9. (Currently amended) The method of claim 7, ~~wherein~~ where the variation of context in which the sequence occurs, $H(S)$, is calculated as

$$HM(S) = MIN(HL(S), HR(S)),$$

$$HLM(S) = -\sum_{w} \frac{f(wS)}{f(S)} \cdot \log\left(\frac{f(wS)}{f(S)}\right),$$

and

$$HR(S) = -\sum_{w} \frac{f(Sw)}{f(S)} \cdot \log\left(\frac{f(Sw)}{f(S)}\right),$$

where *MIN* defines a minimum operation, *S* represents the sequence, *f(wS)* defines a number of times a particular term, *w,* appears in the collection of documents followed by the sequence, *f(Sw)* refers to a number of times the sequence is followed by *w* in the collection of documents, and *f(S)* refers to a number of times the sequence *S* is present in the collection of documents.

10. (Currently amended) The method of claim 7, ~~wherein~~ where the variation of context in which the sequence occurs, *HM(S),* is calculated as

$$HM(S) = MIN(HLM(S), HRM(S)),$$

where MIN defines a minimum operation, HLM*(S)* is defined as a minimum of $1 - \frac{f(wS)}{f(S)}$ for each term w in the collection of documents, *HRM(S)* is defined as a minimum of $1 - \frac{f(Sw)}{f(S)}$ for each term *w* in the collection of documents, *f(wS)* defines a number of times a particular term, *w,* appears in the collection of documents followed by the sequence, *f(Sw)* refers to a number of times the sequence is followed by *w* in the collection of documents, and *f(S)* refers to a number of times the sequence is present in the collection of documents.

11. (Currently amended) The method of claim 7, ~~wherein~~ where the variation of context in which the sequence occurs, *HC(S),* is calculated as

$$HC(S) = MIN(HLC(S), HRC(S)),$$

where MIN defines a minimum operation, *HLC(S)* is defined as $\sum_{w} \delta(wS)$ and *HRC(S)* is defined as $\sum_{w} \delta(Sw)$, where $\delta(X)$ is defined as one if sequence X occurs in the collection of documents and zero otherwise, where wS refers to a particular word followed by the sequence, and where Sw refers to the sequence followed by a word.

12. (Currently amended) The method of claim 7, ~~wherein~~ where the variation of context in which the sequence occurs, *HP(S),* is calculated as

$$HP(S) = MIN(HLP(S), HRP(S))$$

where MIN defines a minimum operation, *HLP(S)* is defined as the number of continuations to the left of the sequence that cover a predetermined percentage of all cases in the collection of documents and *HRP(S)* is defined as the number of continuations to the right of the sequence that cover the predetermined percentage of all cases in the collection of documents.

13. (Canceled)

14. (Currently amended) The method of claim 1, ~~wherein~~ where the sequence includes three or more words.

15. (Currently amended) The method of claim 1, further including:

applying one or more rules to the sequence, and

~~wherein~~ where ~~determining whether~~ identifying that the sequence is a semantic unit is further based at least in part on the application of the one or more rules.

16. (Currently amended) A device comprising:

a memory to store instructions; and

a processor to execute the instructions to implement:

a receiving component ~~configured~~ to receive a sequence of terms;

a coherence component ~~configured~~ to calculate a coherence of multiple terms in the sequence of terms;

a variation component ~~configured~~ to calculate a variation of context terms in a collection of documents in which the sequence occurs, where the variation of context terms is calculated as a measure of entropy of the context of the sequence; and

a decision component ~~configured~~ to determine whether the sequence constitutes a semantic unit based at least in part on results of the coherence component and the variation component, and output an indication of whether the sequence constitutes a semantic unit for use in a processor.

17. (Currently amended) The device of claim 16, ~~wherein~~ where the context terms include terms to the left and right of the sequence.

18. (Currently amended) The device of claim 16, ~~wherein~~ where the coherence of the terms in the sequence is calculated relative to the collection of documents.

19. (Currently amended) The device of claim 18, ~~wherein~~ where the coherence of the terms in the sequence is calculated as a likelihood ratio that defines a probability of the sequence occurring in the collection of documents relative to parts of the sequence occurring.

20. (Canceled)

21. (Currently amended) The device of claim [[20]] 16, ~~wherein~~ where the variation of context in which the sequence occurs, *H(S)*, is calculated as

$$H(S) = MIN\big(HL(S), HR(S)\big),$$

$$HL(S) = -\sum_{w} \frac{f(wS)}{f(S)} \cdot \log\left(\frac{f(wS)}{f(S)}\right),$$

and

$$HR(S) = -\sum_{w} \frac{f(Sw)}{f(S)} \cdot \log\left(\frac{f(Sw)}{f(S)}\right),$$

where *MIN* defines a minimum operation, *S* represents the sequence, *f(wS)* defines a number of times a particular term, *w*, appears in the collection of documents followed by the sequence, *f(Sw)* refers to a number of times the sequence is followed by *w* in the

collection of documents, and *f(S)* refers to a number of times the sequence *S* is present in the collection of documents.

22. (Currently amended) The device of claim [[20]] 16, ~~wherein~~ where the variation of context in which the sequence occurs, *HM(S),* is calculated as

$$HM(S) = MAX(HLM(S), HRM(S)),$$

where MIN defines a minimum operation, HLM*(S)* is defined as a minimum of $1 - \frac{f(wS)}{f(S)}$ for each term w in the collection of documents, *HRM(S)* is defined as a minimum of $1 - \frac{f(Sw)}{f(S)}$ for each term *w* in the collection of documents, *f(wS)* defines a number of times a particular term, *w,* appears in the collection of documents followed by the sequence, *f(Sw)* refers to a number of times the sequence is followed by *w* in the collection of documents, and *f(S)* refers to a number of times the sequence is present in the collection of documents.

23. (Currently amended) The device of claim [[20]] 16, ~~wherein~~ where the variation of context in which the sequence occurs, *HC(S),* is calculated as

$$HC(S) = MIN(HLC(S), HRC(S)),$$

where MIN defines a minimum operation, *HLC(S)* is defined as $\sum_{w} \delta(wS)$ and *HRC(S)* is defined as $\sum_{w} \delta(Sw)$, where $\delta(X)$ is defined as one if sequence *X* occurs in the

document collection and zero otherwise, where wS refers to a particular word followed by the sequence, and where Sw refers to the sequence followed by a word.

24. (Currently amended) The device of claim [[20]] 16, ~~wherein~~ where the variation of context in which the sequence occurs, *HP(S),* is calculated as

$$HP(S) = MIN(HLP(S), HRP(S))$$

where MIN defines a minimum operation, *HLP(S)* is defined as the number of continuations to the left of the sequence that cover a predetermined percentage of all cases in the collection of documents and *HRP(S)* is defined as the number of continuations to the right of the sequence that cover the predetermined percentage of all cases in the collection of documents.

25. (Currently amended) The device of claim 16, ~~wherein~~ where the decision component is further ~~configured~~ to compare the results of the coherence component and the variation component to threshold values and identify the sequence as a semantic unit based at least in part on the comparisons.

26. (Currently amended) The device of claim 16, ~~further comprising~~ where the processor further executes the instructions to implement:

a heuristics component ~~configured~~ to apply one or more predefined rules to the sequence, ~~wherein~~ where the decision component is further ~~configured~~ to determine whether the sequence constitutes a semantic unit based at least in part on application of the one or more rules.

27. (Currently amended) The device of claim 26, ~~wherein~~ where the one or more rules are exclusionary rules that determine when certain sequences are not semantic units.

28. (Currently amended) A device comprising:

a memory to store instructions; and

a processor to execute the instructions to implement:

means for receiving a sequence of terms;

means for calculating a first value representing a coherence of terms in the sequence of terms;

means for calculating a second value representing variation of context in which the sequence occurs;

means for comparing the first value to a first threshold and the second value to a second threshold;

means for ~~determining whether~~ identifying that the sequence is a semantic unit based at least in part on the first value satisfying the first threshold and second ~~values~~ value satisfying the second threshold; and

means for outputting an indication ~~of whether~~ that the sequence is a semantic unit based on identifying that the sequence is a semantic unit ~~for use in a processor~~.

29. (Currently amended) A computer-readable memory device that includes programming instructions ~~configured~~ to control at least one processor, the computer-readable memory device comprising:

instructions for calculating a first value representing a coherence of terms in a sequence of terms;

instructions for calculating a second value representing variation of context in which the sequence occurs, where the variation of context in which the sequence occurs is calculated as a measure of entropy of the context of the sequence;

instructions for ~~determining whether~~ identifying that the sequence is a semantic unit based on the first and second values; and

instructions for outputting an indication ~~of whether~~ that the sequence is a semantic unit.

30. (Currently amended) The computer-readable memory device of claim 29, ~~wherein~~ where the coherence of the terms in the sequence is calculated relative to a collection of documents.

31. (Currently amended) The computer-readable memory device of claim 30, ~~wherein~~ where the coherence of the terms in the sequence is calculated as a likelihood ratio that defines a probability of the sequence occurring in the collection of documents relative to parts of the sequence occurring.

32. (Currently amended) The computer-readable memory device of claim 30, ~~wherein~~ where the coherence of the terms in the sequence is calculated as:

$$LR(A,B) = \frac{L(f(B),N)}{L(f(AB),f(A)) \cdot L(f(\sim AB),f(\sim A))},$$

where $f(A)$ defines a number of occurrences of term A in the collection of documents, $f(\sim A)$ defines a number of occurrences of a term other than term A in the collection of documents, $f(B)$ defines a number of occurrences of term B in the collection of documents, N defines a total number of events in the collection of documents, $f(AB)$ defines a number of times term A is followed by term B in the collection of documents, and $f(\sim AB)$ is a number of times a term other than A is followed by term B in the collection of documents, ~~wherein~~ where

$$L(k,n) = \left(\frac{k}{n}\right)^{k} \cdot \left(1-\frac{k}{n}\right)^{(n-k)}.$$

33. (Currently amended) The computer-readable memory device of claim 29, ~~wherein~~ where the coherence of the terms in the sequence are defined as not being sufficient unless a threshold is met.

34. (Currently amended) The computer-readable memory device of claim 33, ~~wherein~~ where the threshold is defined as: $f(AB) > \frac{f(A) \cdot f(B)}{N}$, where $f(A)$ defines a

number of occurrences of term *A* in the collection of documents, *f(B)* defines a number of occurrences of term *B* in the collection of documents, *N* defines a total number of events in the collection of documents, and *f(AB)* defines a number of times term A is followed by term *B* in the collection of documents.

35. (Currently amended) The computer-readable memory device of claim 29, ~~wherein~~ where the variation of context in which the sequence occurs is calculated relative to a collection of documents.

36. (Canceled)

37. (Currently amended) The computer-readable memory device of claim 35, ~~wherein~~ where the variation of context in which the sequence occurs, *H(S)*, is calculated as

$$HM(S) = MIN\left(HL(S), HR(S)\right),$$

$$HLM(S) = -\sum_{w} \frac{f(wS)}{f(S)} \cdot \log\left(\frac{f(wS)}{f(S)}\right),$$

and

$$HR(S) = -\sum_{w} \frac{f(Sw)}{f(S)} \cdot \log\left(\frac{f(Sw)}{f(S)}\right),$$

where *MIN* defines a minimum operation, *S* represents the sequence, *f(wS)* defines a number of times a particular term, *w*, appears in the collection of documents followed by the sequence, *f(Sw)* refers to a number of times the sequence is followed by *w* in the

collection of documents, and *f(S)* refers to a number of times the sequence *S* is present in the collection of documents.

38. (Currently amended) The computer-readable memory device of claim 35, ~~wherein~~ where the variation of context in which the sequence occurs, *HM(S),* is calculated as

$$HM(S) = MIN(HLM(S), HRM(S)),$$

where MIN defines a minimum operation, HLM*(S)* is defined as a minimum of $1 - \frac{f(wS)}{f(S)}$ for each term w in the collection of documents, *HRM(S)* is defined as a minimum of $1 - \frac{f(Sw)}{f(S)}$ for each term *w* in the collection of documents, *f(wS)* defines a number of times a particular term, *w,* appears in the collection of documents followed by the sequence, *f(Sw)* refers to a number of times the sequence is followed by *w* in the collection of documents, and *f(S)* refers to a number of times the sequence is present in the collection of documents.

39. (Currently amended) The computer-readable memory device of claim 35, ~~wherein~~ where the variation of context in which the sequence occurs, *HC(S),* is calculated as

$$HC(S) = MIN(HLC(S), HRC(S)),$$

where MIN defines a minimum operation, *HLC(S)* is defined as $\sum_{w} \delta(wS)$ and *HRC(S)* is defined as $\sum_{w} \delta(Sw)$, where $\delta(X)$ is defined as one if sequence X occurs in the collection of documents and zero otherwise, where wS refers to a particular word followed by the sequence, and where Sw refers to the sequence followed by a word.

40. (Currently amended) The computer-readable memory device of claim 35, ~~wherein~~ where the variation of context in which the sequence occurs, *HP(S),* is calculated as

$$HP(S) = MIN(HLP(S), HRP(S))$$

where MIN defines a minimum operation, *HLP(S)* is defined as the number of continuations to the left of the sequence that cover a predetermined percentage of all cases in the collection of documents and *HRP(S)* is defined as the number of continuations to the right of the sequence that cover the predetermined percentage of all cases in the collection of documents.

41. (Currently amended) The computer-readable memory device of claim 29, ~~wherein~~ where the instructions for ~~determining whether~~ identifying that the sequence is a semantic unit include instructions for comparing the first and second values to first and second thresholds and identifying the sequence as a semantic unit when the first and second values satisfy the first and second thresholds.

42. (Currently amended) The computer-readable memory device of claim 29, ~~wherein~~ where the sequence includes three or more words.

43. (Currently amended) The computer-readable memory device of claim 29, further including:

instructions for applying one or more rules to the sequence, and

~~wherein~~ where the instructions for ~~determining whether~~ identifying that the sequence is a semantic unit are further based at least in part on the application of the one or more rules.

44. (New) The system of claim 28, where the coherence of the terms in the sequence is calculated as a likelihood ratio that defines a probability of the sequence occurring in the collection of documents relative to parts of the sequence occurring.

45. (New) The system of claim 28, where the variation of context in which the sequence occurs is calculated as a measure of entropy of the context of the sequence.